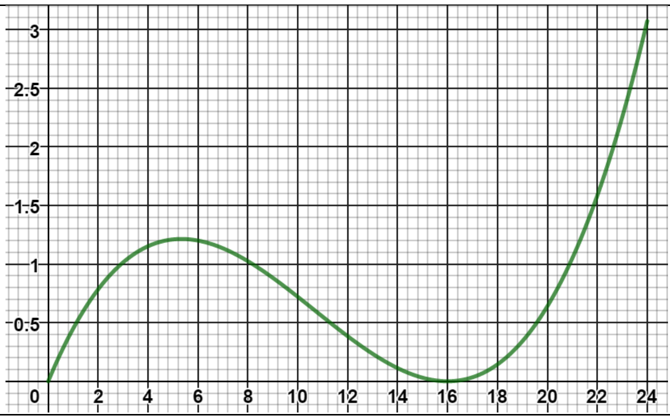
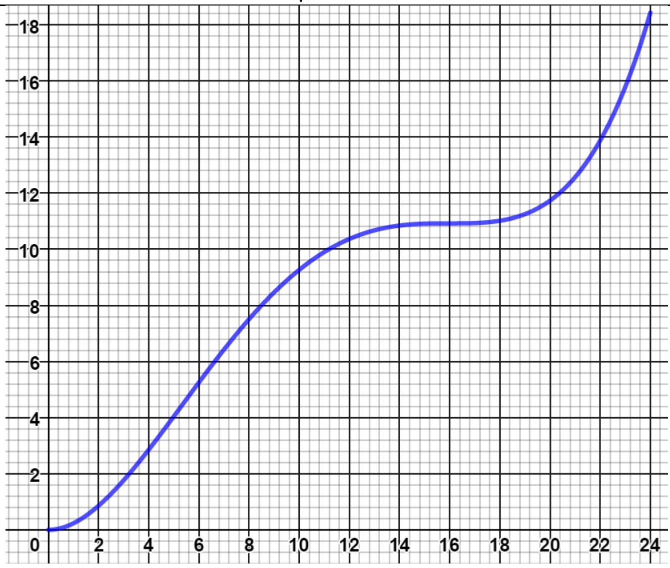
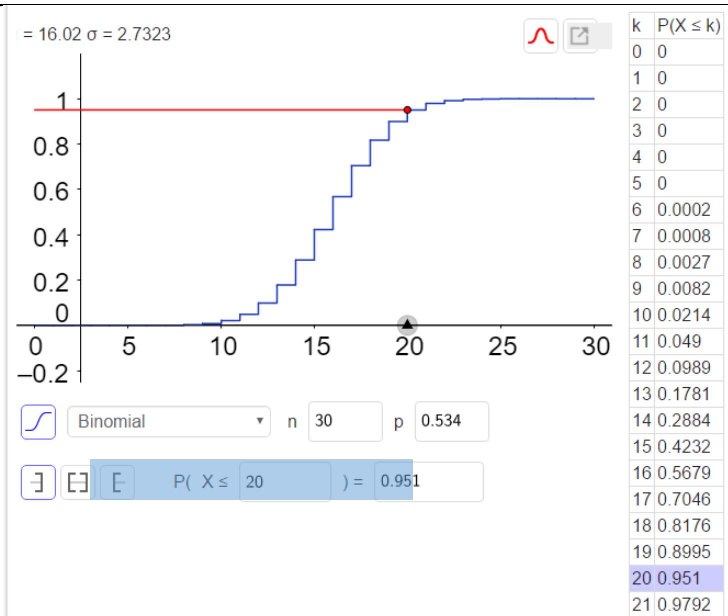
Réponses de la partie B

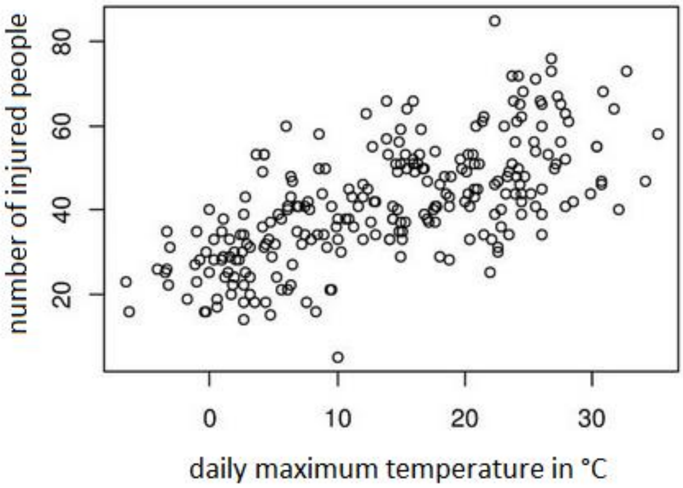
	Partie B1	Points			
		CC	M	RP	I
Les températures moyennes par mois ont été enregistrées en 2002 au Grand-Duché de Luxembourg. On sait que janvier 2002 a été le mois le plus froid avec 1,6°C et que la température moyenne la plus élevée a été mesurée en juin 2002 avec 18,6°C.					
a)	**Justifier**, qu'en Europe, les températures moyennes mensuelles pour quelques années consécutives peuvent être modélisées avec un modèle périodique.				
	Les températures augmentent et diminuent au cours d'une année et cette procédure se répète chaque année. Bien que les températures ne soient pas les mêmes chaque année, en moyenne les mois d'été sont chauds et les mois d'hiver sont froids.	2			
b)	**Donner** l'amplitude et la période de ce modèle.				
	Amplitude $a = 8{,}5$ Période $T = 12$	1	1		
c)	**Déterminer les** paramètres a, b, c et d dans le modèle du type $T(x) = a \cdot sin(b \cdot (x - c)) + d$ qui décrit les données où T est la température moyenne et x est le mois, en commençant par $x = 1$ pour janvier 2002.				
	$a = \frac{18{,}6 - 1{,}6}{2} = 8{,}5$ $d = 8{,}5 + 1{,}6 = 10{,}1$ $b = \frac{2\pi}{12} \approx 0{,}524$ $c = 4$ parce que $\text{Extremum}\left(\frac{17}{2} \sin\left(\frac{131}{250} x\right) + \frac{101}{10}, -4, -2\right)$ $\rightarrow \mathbf{(-3, 1.6)}$ il y a donc un minimum en $x = -3$ qui doit être déplacé vers $x = 1$.	1	1 1	1 1	
Les précipitations ont été observées un jour précis en mars 2002. Les précipitations de ce jour peuvent être modélisées par la fonction $R(t) = 0.002t^3 - 0.064t^2 + 0.512t, 0 \leq t \leq 24$ où $R(t)$ est le taux de précipitation en mm/h et t est le temps en heures.					
d)	**Décrire**, à l'aide d'un court texte descriptif, cette journée en termes de précipitations. Votre réponse doit se concentrer sur les moments où il pleut le plus et le moins.				
	Il commence à pleuvoir à minuit et les précipitations atteignent leur maximum entre 5 et 6 heures. La pluie diminue jusqu'à 16 heures, où elle s'arrête un peu. Dans l'après-midi, la pluie augmente à nouveau jusqu'à atteindre un maximum à minuit.			2	1

					
Un cylindre de verre vide a été placé à l’extérieur pendant cette journée pour aider à voir combien de pluie était tombée.					
e)	**Esquisser** le graphique d’une fonction qui montre la hauteur de l’eau dans un cylindre de verre, qui était placé à l’extérieur et vide auparavant.				
	 Le graphique montre la fonction dérivée de $R(t)$.	1	1	1	
f)	**Calculer** la quantité totale de pluie de ce jour en mm.				
	$\int_0^{24} R(t) = 18{,}43$ Il est tombé environ 18,4 mm de pluie ce jour-là.		1	1	
L’année 2002 a connu 195 jours de pluie et 170 jours sans pluie au Grand-Duché de Luxembourg. On peut supposer que tous les jours ont la même chance d’être un jour de pluie. Un an plus tard, les météorologues veulent savoir s’il y a eu plus de pluie en 2003. Malheureusement, certaines données ont été perdues, et ils n’ont donc pris qu’un petit échantillon de 30 jours consécutifs.					
g)	**Calculer** la probabilité qu’il pleuve un jour au hasard, si l’on suppose que le nombre total de jours de pluie dans les deux années reste constant et que les jours de pluie sont également répartis sur toute l’année.				
	$p = \frac{195}{365} \approx 0{,}534$	1			
h)	**Utiliser** un test d’hypothèse nulle pour déterminer combien de jours il doit pleuvoir pour que les météorologues puissent dire qu’il y a eu plus de pluie en 2003 qu’en 2002 lorsque le seuil de signification est de 5%.				
	$H_0 : p > 0{,}534$ $H_1 : p \leq 0{,}534$		2		



	$k = 20$ Ainsi, s'il a plu pendant plus de 20 jours, les scientifiques peuvent être sûrs à 95 % qu'il a plu davantage en 2003 qu'en 2002.		1	1	1

Le diagramme suivant montre la température maximale et le nombre de blessés causés par les accidents de la circulation à Berlin à long terme.



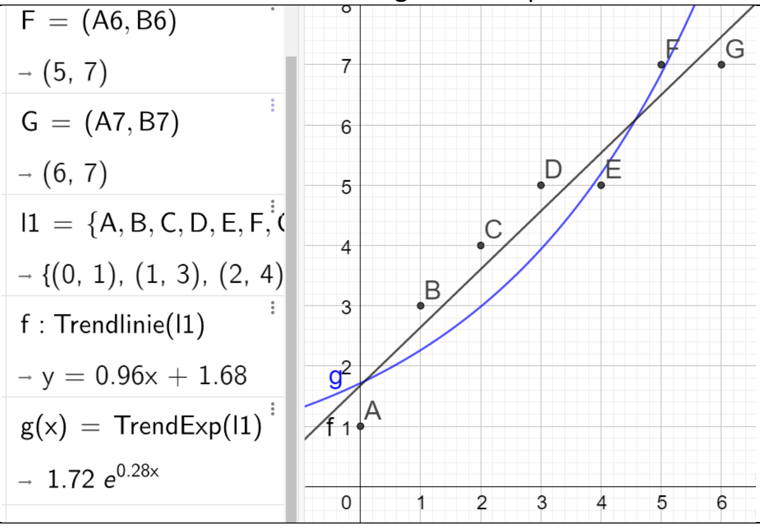
i)	**Décrire** la corrélation entre ces deux valeurs.				
	Plus la température est élevée, plus le nombre de blessés est important.				1
j)	**Expliquer** pourquoi le nombre de blessés peut être corrélé de cette manière avec la température maximale.				
	Lorsqu'il fait chaud, plus de gens sortent et le risque d'être impliqué dans un accident de voiture est donc plus élevé que si les gens restent chez eux.				1
	Total = 25	**6**	**8**	**7**	**4**

	Partie B2	Points			
		CC	M	RP	I
Dans une station de test Covid-19, 19 personnes présentant des symptômes ont été testées un jour donné et 6 d'entre elles ont eu un résultat positif. Le même jour, 87 personnes sans symptômes ont été testées, dont 85 ont eu un résultat négatif.					
a)	**Montrer** que la probabilité d'obtenir un résultat positif dépend du fait qu'une personne présente ou non des symptômes.				
	Considérons les événements suivants : S = « la personne a des symptômes » P = « la personne est positive » $P(P) = \frac{8}{106}$ et $P(P\|S) = \frac{6}{19}$. $P(P) \neq P(P\|S)$, donc la probabilité d'obtenir un résultat positif dépend du fait qu'une personne présente ou non des symptômes.		2		
Pour protéger les données personnelles, les enquêtes de test sont étiquetées avec un code, qui contient 2 lettres (parmi un alphabet de 26 lettres) et 4 chiffres (0-9). Les mêmes lettres et chiffres peuvent être choisis plus d'une fois.					
b)	**Calculer** le nombre total de codes différents pouvant être créés par ce système.				
	Le nombre total de codes différents est $26^2 \cdot 10^4 = 6\,760\,000$.	1	1		
Après quelques mois, les statistiques ont montré que 1,7% des personnes sans symptômes sont testées positives. Une entreprise comptant 20 employés (tous sans symptômes) demande à tous ses employés de se faire tester.					
c)	**Donner** deux hypothèses, qui doivent être faites pour modéliser cette situation avec une distribution binomiale.				
	• Pour la distribution binomiale, tous les essais doivent être indépendants. Nous devons donc supposer que les employés n'ont aucune possibilité de se contaminer entre eux. • Nous savons également que certaines personnes (certains groupes d'âge ou problèmes de santé) sont plus susceptibles d'être infectées. Nous devons donc supposer que le risque d'infection est le même pour tous les employés. • De plus, nous devons supposer que la population de la statistique constitue un très grand nombre de personnes. • Nous devons supposer que les tests sont fiables. • Nous devons supposer qu'aucun employé n'a été vacciné. • Nous devons supposer qu'aucun des employés n'a eu la Covid-19 auparavant, de sorte qu'il pourrait être immunisé maintenant.	1	1		
d)	**Calculer** la probabilité qu'au moins un des employés soit testé positif.				
	Calcul de la probabilité qu'aucun des employés ne soit testé positif : $P(X=0) = B_{200;0{,}017}(0) = \binom{200}{0} \cdot 0{,}017^0 \cdot 0{,}983^{200}$ Donc $P(X \geq 1) = 1 - P(X=0) = 1 - 0{,}017^0 \cdot 0{,}983^{200}$ ce qui donne $P(X \geq 1) \approx 0{,}968$		2	1	
En supposant que la situation puisse être modélisée par une distribution binomiale donnée par la formule $B(84; 0{,}02; k) = \binom{84}{k} \cdot 0{,}02^k \cdot 0{,}98^{84-k}$.					
e)	**Interpréter les** valeurs 84, 0,02 et 0,98 dans le contexte donné.				
	84 est le nombre total d'employés, 0,02 est la probabilité qu'un employé soit testé positif et 0,98 est la probabilité de l'événement contraire, c'est-à-dire la probabilité qu'un employé ne soit pas testé positif.	1	2		
Le 5 mars 2020, un homme de retour d'Italie est la première personne au Grand-Duché de Luxembourg à avoir été testée positive au COVID-19. Ce jour est donc marqué comme le jour 0 dans la statistique. Le					

tableau suivant montre le nombre total de personnes infectées enregistrées au Luxembourg dans les jours qui ont suivi l'apparition du premier cas.

Jour	0	1	2	3	4	5	6
Nombre	1	3	4	5	5	7	7

f) **Dessiner** le nuage de points (graphique de dispersion) de ces valeurs et ainsi qu'un modèle de régression linéaire et un modèle de régression exponentiel.



2 | 1

g) **Donner** les équations qui décrivent les lignes de régression de la question f).

Modèle linéaire : $f(x) = 0{,}96x + 1{,}68$
Modèle exponentiel : $1{,}72 \cdot e^{0{,}28x}$

2

h) **Expliquer** pourquoi il était si difficile de décider si la propagation du virus est mieux modélisée par un modèle linéaire ou exponentiel à ce stade précoce.

À un stade précoce, les deux régressions semblent bien adaptées aux données. Les valeurs aberrantes ne compromettent pas de manière critique l'un ou l'autre de ces modèles. Il y a peut-être une valeur aberrante supplémentaire pour le modèle exponentiel, mais la courbe passe par deux points alors qu'elle touche à peine deux points pour le modèle linéaire.

2

Après sept jours supplémentaires, d'autres modèles ont été réalisés pour faire de meilleures prédictions, où t est donné en jours :

$$A(t) = 1.35567 \cdot 1.46977^{t}$$
$$B(t) = 12.4396 \cdot t - 34.8571$$

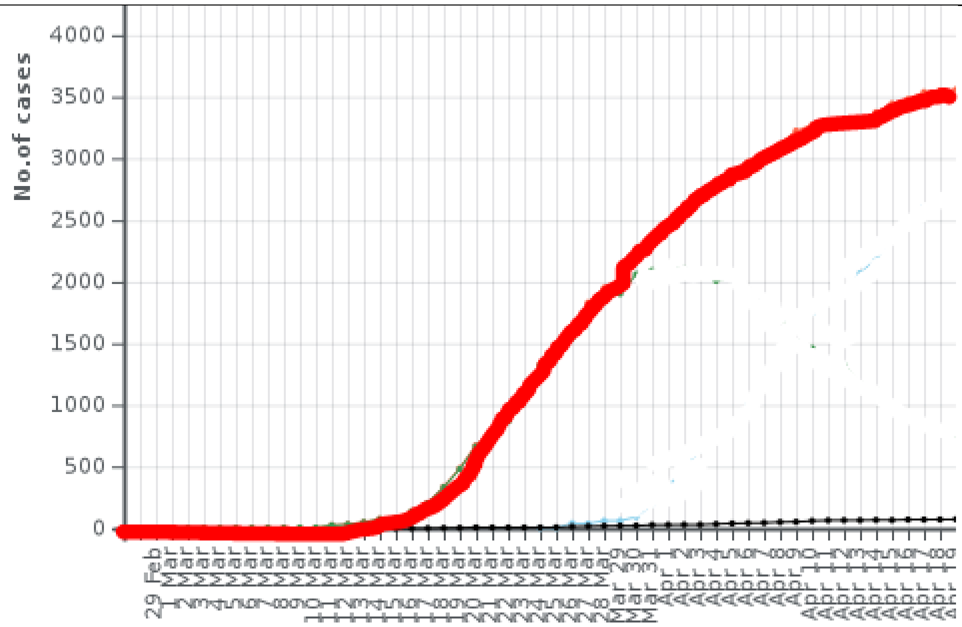
Au 16ème jour, 670 cas de COVID-19 ont été enregistrés au Luxembourg.

i) **Calculer** le nombre prévisionnel de personnes infectées le 16ème jour avec le modèle A, puis avec le modèle B et **comparer** ces nombres avec le nombre réellement observé. **Décider** du modèle le mieux adapté à cette situation et **justifier** la réponse.

$A(16) = 1.35567 \cdot 1.46977^{16}$ donc $A(16) \approx 643$
$B(16) = 12.4396 \cdot t - 34{,}8571$ donc $B(16) \approx 164$
La valeur prédite pour le modèle A est proche de la valeur observée. La valeur prédite pour le modèle B est beaucoup plus petite que la valeur observée. Etant donné le grand écart entre les deux valeurs prédites, le modèle A est plus adapté à la situation.

1 | 1

Le diagramme suivant montre l'évolution du nombre total d'infections enregistrées pour les 4 premières semaines au Grand-Duché de Luxembourg.

					
j)	**Donner** deux raisons possibles pour lesquelles la courbe s'aplatit à un stade ultérieur.				
	• (En utilisant l'outil technologique) nous pouvons voir que $\lim_{t\to\infty} C(t) = 3404$ ce qui montre que la courbe a une tangente horizontale et s'aplatit donc lorsque t augmente. • $\lim_{t\to\infty} C'(t) = 0$ ce qui montre qu'au fur et à mesure que t augmente, la tangente à la courbe se déplace vers une position horizontale et donc la courbe s'aplatit. Des raisons non mathématiques peuvent également être données : • En été, les taux d'infection diminuaient, car beaucoup de gens sortent. • Les magasins et les écoles ont été fermés, ce qui a entraîné une baisse du taux d'infection. • Les gens portaient plus de masques. D'autres raisons qui pourraient être fausses, mais qui expliqueraient quand même le résultat : • Les personnes qui ont eu la Covid-19 sont immunisées, donc la population susceptible d'être contaminée diminue. • Le nombre d'essais a diminué.			1	1
La courbe peut être modélisée par la fonction $$C(t) = \frac{3404}{1 + 193 \cdot e^{-0.233 \cdot t}}$$					
k)	**Déterminer** le jour où le taux d'infection est le plus élevé.				
	Nous devons trouver la valeur maximale de la dérivée sur l'intervalle [0 ; 28]. En utilisant l'outil technologique, nous trouvons que la valeur maximale de C' est atteinte pour $t \approx 22{,}59$, c'est-à-dire que le 23$^{\text{ème}}$ jour, le taux d'infection a atteint son maximum.	1	1		
	Total = 25	**7**	**13**	**4**	**1**